PB84-217777

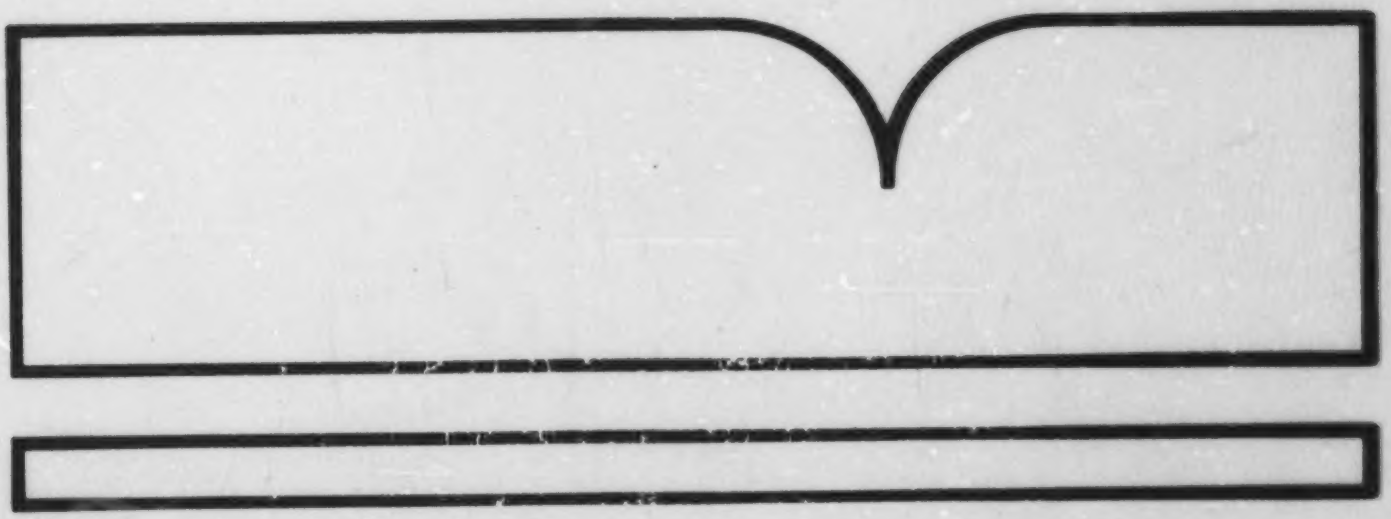
Chemical Information System
A Blue Ribbon Panel Review

Life Systems, Inc., Cleveland, OH

Prepared for

Environmental Protection Agency, Washington, DC

Aug 84



U.S. Department of Commerce
National Technical Information Service
NTIS

United States
Environmental Protection
Agency

Office of Information
Resources Management
Washington DC 20460

August 1984
EPA 200-02-84-002
PB 84-217777

EPA

Chemical Information System:

A Blue Ribbon Panel Review

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA 200-02-84-002		PB 84-217777

4. TITLE AND SUBTITLE	5. REPORT DATE
Chemical Information System: A Blue Ribbon Panel Review	August 1984
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Life Systems, Inc. 24755 Highpoint Road Cleveland, Ohio 44122	10. PROGRAM ELEMENT NO.
	11. CONTRACT/GRANT NO. 68-02-4038
12. SPONSORING AGENCY NAME AND ADDRESS Office of Information Resources Management U.S. Environmental Protection Agency 401 M St., S.W. Washington, D.C. 20460	13. TYPE OF REPORT AND PERIOD COVERED
	14. SPONSORING AGENCY CODE EPA 400/03

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The Chemical Information System (CIS) was initiated in 1971 by the Environmental Protection Agency and the National Institutes of Health. The CIS provides chemical information through computerized data bases to Government agencies and the public. At the request of the EPA Office of Information Resources Management, Interdisciplinary Consulting and Information Research of Life Systems, Inc. assembled a panel of nine experts to review CIS. The panel met at a symposium on June 11-13, 1984 to make recommendations for technical, financial and managerial improvements for CIS.

This Symposium Summary Report presents the consensus achieved during the symposium.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
NTIS, Springfield, VA 22161	UNCLASSIFIED	29
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

CHEMICAL INFORMATION SYSTEM:

A BLUE RIBBON PANEL REVIEW

Prepared under Contract No. 68-02-4038
for the U.S. Environmental Protection Agency
Office of Information Resources Management
by Life Systems, Inc.

August 1984
EPA 200-02-84-002
PB 84-217777

TABLE OF CONTENTS

		PAGE
LIST OF TABLES		ii
LIST OF ACRONYMS		iii
1.0	EXECUTIVE SUMMARY	1-1
2.0	INTRODUCTION	2-1
2.1	Background	2-1
2.2	Objective	2-1
2.3	Approach	2-1
3.0	CONSENSUS STATEMENTS	3-1
3.1	CIS Purpose	3-1
3.2	CIS Capabilities and Characteristics: Existing Versus Proposed	3-1
3.3	Specific CIS Management Problems	3-4
3.4	Steps Toward Becoming Self-Supporting	3-4
3.5	Data Bases or Files Which Could Now be Self-Supporting	3-8
3.6	Government Responsibility	3-9
3.7	Appropriate CIS Management Structure	3-9
3.7.1	Responsibilities	3-9
3.7.2	Advantages and Disadvantages for Potential Implementing Organizations	3-10
3.7.3	Recommendation for Implementing Organization	3-12
3.7.4	Appropriate Management Organization	3-12
3.7.5	Summary	3-13
APPENDIX		
1	List of Participants	A1-1
2	List of Background References Supplied to Panelists	A2-1
3	Agenda for CIS Symposium	A3-1

LIST OF TABLES

TABLE		PAGE
3-1	Comparison of Proposed with Existing CIS Capabilities and Characteristics	3-2
3-2	Possible Causes of Inadequate CIS Management Performance .	3-5
3-3	Examples of CIS Problems by Decision-Making Area	3-6

LIST OF ACRONYMS

ACS	American Chemical Society
CAS	Chemical Abstract Service
CEQ	Council on Environmental Quality
CMA	Chemical Manufacturer's Association
CIS	Chemical Information System
CPSC	Consumer Product Safety Commission
CTCP	Clinical Toxicology of Commerical Products
DOE	Department of Energy
DOE/RECON	Remote Console (Information Retrieval System Operated by DOE)
DOT/CG	Department of Transportation/Coast Guard
EPA	Environmental Protection Agency
FASEB	Federation of American Societies for Experimental Biology
FDA	Food and Drug Administration
FRSS	Federal Register Search System
ICAIR	Interdisciplinary Consulting and Information Research
MSSS	Mass Spectral Search System
NBS	National Bureau of Standards
NIH	National Institutes of Health
NIOSH	National Institute for Occupational Safety and Health
NLM	National Library of Medicine
OHMTADS	Oil and Hazardous Material Technical Assistance Data System
OIRM	Office of Information Resources Management
RTECS	Registry of Toxic Effects of Chemical Substances
SANSS	Structure and Nomenclature Search System
SPHERE	Scientific Parameters in Health and the Environment, Retrieval and Estimation
SRI	Stanford Research Institute
TSCAPP	Toxic Substances Control Act Plant and Production

1.0 EXECUTIVE SUMMARY

The Chemical Information System (CIS) was initiated in 1971 by the Environmental Protection Agency (EPA) and the National Institutes of Health (NIH). The CIS provides chemical information through computerized data bases to government agencies and the public. At the request of the EPA Office of Information Resources Management (OIRM), ICAIR (Interdisciplinary Consulting and Information Research) of Life Systems, Inc. assembled a panel of nine experts to review CIS. The panel met at a symposium on June 11-13, 1984 to make recommendations for technical, financial and managerial improvements for CIS. The panel included experts in the following fields: chemistry, toxicology, numerical data, environmental and information science.

Overall, the panel concluded that CIS as a concept has considerable inherent value, however, improvements are necessary before CIS could serve the significant need for a high quality, integrated, scientific and technical chemical information system.

The panel reached unanimity of opinion on seven major topics: CIS purpose, capabilities, management problems, steps to become self-supporting, potential self-supporting files, government responsibilities and the management structure most appropriate for CIS. Briefly, the specific recommendations or consensus statements developed by the panel are as follows:

- After reviewing the demands for chemical information, the panel concluded that "...there is a need for a high quality, integrated scientific and technical system that is widely and easily accessible with manipulation and modeling capabilities."

  The panel was informed by EPA that the original purpose of CIS (to provide mass spectral data) had never been formally revised, and that a more appropriate purpose was needed for the current system. Therefore, the panel proposed that "the purpose of CIS is to be an integrated system of chemical data (including both qualitative and quantitative data) which serves the needs of governmental agencies and the public for data on environmental and health effects and characteristic properties of chemical substances." The panel recommended that this purpose be adopted and serve as the focus for a new management strategy.

- The panel concluded that existing CIS capabilities and characteristics are inadequate to meet what they perceived are the demands for chemical information. Lack of consistency in mnemonics, terminology, educational and training materials highlight a list of deficiencies which reduce the effectiveness of CIS. However, the panel noted that the directory capability achieved by the Structure and Nomenclature Search System (SANSS) and the numeric data bases in CIS are exceptional.

- The panel judged that CIS data base selection, implementation and maintenance, systems standards, quality control, policy-making, management resources and public awareness are inadequate. Data base selection, implementation and public awareness of CIS require the most improvement.

- The panel concluded that CIS should not be required to become entirely self-supporting. Underutilized, but valuable data bases will need continued government support. However, many steps can and should be taken to make CIS more self-supporting than at present. These steps include: retrieval of more costs from users, improvement in operating efficiency, and increased utilization of the system.

- Although the panel agreed that all high use files presently on CIS could now be self-supporting if transferred to the private sector, such fragmentation of CIS as a system would have a detrimental impact to the ascribed purpose of providing an integrated system. As long as CIS can be improved to meet the specifications developed by the panel, high use files should not be separated. This does not preclude concurrent availability of high use files in the private sector which might be a source of additional revenue to CIS.

- In the panel's view, the government must continue to be responsible for assuring the existence of a publicly available system supplying chemical information. This responsibility should be implemented through a new management strategy capable of correcting the major problems and enhancing the ability of CIS to serve its purpose.

- The panel concluded with a recommendation that a more appropriate management concept for CIS be adopted. The panel's recommendation for this concept consists of a management role for the government and a separate implementation role. For the management role within the government the panel recommended EPA, the National Library of Medicine (NLM) or the Council on Environmental Quality (CEQ), in that order of preference. The panel recommended NLM, a professional society, or a private data base vendor as viable candidates for the implementation role. As second choices, but also viable candidates, the panel recommended a privately operated government laboratory or a private research institute.

2.0 INTRODUCTION

2.1 Background

In 1970, NIH initiated a computerized system to provide mass spectral data. Subsequently, EPA and NIH collaborated to add other data bases and create CIS containing a variety of chemical data and information. Other Federal Agencies such as the National Bureau of Standards (NBS), the Food and Drug Administration (FDA), the National Institute for Occupational Safety and Health (NIOSH), the Consumer Product Safety Commission (CPSC) and the Department of Transportation-Coast Guard (DOT/CG) assisted in the continuing development of CIS by contributing data bases. Offering integrated files of wide ranging chemical information, CIS is available to the public through local telephone lines through telecommunication networks and is now used by over 600 organizations in 20 countries.

Continued growth in size, cost and complexity has led to problems in CIS operation and maintenance. Current EPA costs to operate and maintain the system exceed $1.6 million per year. At the end of 1983, user service was temporarily halted due to lack of funds. An Inspector General's Office probe into financial management of CIS has ensued. The computer system used to store and operate CIS data bases is more than 10 years old, exceeding the normal life cycle of a major computer system. The adequacy of the system is, therefore, in question. Finally, an undetermined amount of the chemical information, formerly unique to CIS, is now available from other sources.

2.2 Objectives

In order to resolve these problems, OIRM sought, as a first step, an independent review of CIS from a panel of chemical information experts (Appendix 1). The objectives of this panel were to review CIS, analyze its weaknesses and strengths, and achieve consensus on recommendations to EPA for technical, financial and management improvements.

2.3 Approach

The OIRM selected a panel symposium approach to obtain the desired review of CIS. Panelists were selected based on their expertise in using data bases, in data base management or in government and private sector chemical information systems. Disciplines represented by the panel included chemistry, toxicology, numerical data, environmental and information science. Panelists included designated representatives from the ACS and the Chemical Manufacturer's Association (CMA). The OIRM staff carefully avoided influencing the panel by contracting the work of selecting, briefing and assembling the panelists to ICAIR. Through ICAIR, OIRM supplied previous CIS reviews by Altman and Fisher (1982) and CRC Systems, Inc. (1983), other background information (Appendix 2) and a CIS user account to each panelist prior to the symposium. Complete financial information on CIS operations, however, was unavailable to the panelists for the symposium.

The OIRM also supplied questions to be answered at the symposium so that panelists could prepare position statements. Panelists worked from a key group of these questions during the symposium. These questions were modified

or appended by the panel in order to address what they considered the primary problem areas of CIS. The symposium agenda (Appendix 3) was also modified to devote adequate time to problem areas considered most important by the panel. The seven major problem areas addressed by the panel were: CIS purpose, capabilities, management problems, steps to become self-supporting, potential self-supporting files, government responsibilities and the management structure appropriate for CIS.

The OIRM staff were present at the symposium only for the opening or to provide additional information at the panel's request.

The symposium was conducted under Work Assignment (WA) 121319 of Contract 68-02-4038. ICAIR of Life Systems, Inc. conducted the symposium which took place on June 11-13, 1984 in Bethesda, Maryland.

This Symposium Summary Report presents the consensus achieved during the symposium. In addition to this introduction, the report includes an executive summary, a description of the panel's consensus statements and recommendations concerning seven major problems areas and the referenced appendices.

3.0 CONSENSUS STATEMENTS AND RECOMMENDATIONS

The panel reached unanimity of opinion on seven major topics. No minority opinions were developed during the Symposium on any of the seven issues.

The panel noted that assessments by Altman and Fisher (1982) and CRC Systems, Inc. (1983) identified similar shortcomings and concluded with many of the same recommendations. A difference between these two reports and the present evaluation is that the need for improvement is now greater.

3.1 CIS Purpose

The OIRM staff, in response to questioning by the panel, indicated that no formal statement of the current purpose of CIS existed. Therefore, before any of the questions of interest to EPA could be addressed, it was necessary to agree upon a purpose and need for CIS. The panel established the following consensus statements as a basis for a need for CIS:

"The demand is growing for increased volumes of reliable, scientific and technical data, and integration and manipulation of data for purposes such as planning, policy making, regulation, research and development.

A number of computer searchable numerical data bases now exist or are under development. Present access systems, however, lack sufficient ease of user access, manipulation capabilities, integration of data bases, modeling capabilities and quality assurance mechanisms."

From these premises, the panel expressed their consensus on the need for CIS in the following statement:

"Therefore, there is a need for a high quality, integrated scientific and technical data system that is widely and easily accessible, with manipulation and modeling capabilities."

The panel developed from this need, a statement of purpose:

"The purpose of CIS is to be an integrated system of chemical data (including both qualitative and quantitative data) to serve the needs of governmental agencies and the public for data on environmental and health effects and characteristic properties of chemical substances.

The CIS should serve these data needs by providing unique numeric data bases, evaluated data, directories of data sources and data manipulation to assist in policy development and regulatory decisions."

The panel used this statement of purpose in reaching consensus on the six remaining topics and recommended that it be adopted and used as a guide by CIS managers.

3.2 CIS Capabilities and Characteristics: Existing Versus Proposed

Table 3-1 lists those capabilities and characteristics recommended by the panel for the chemical information system of the future. This table also presents a comparison between the panelists' perceptions of the existing capabilities and characteristics of CIS and those proposed by the panel for the ideal system. The numerical data on the table indicate cumulative results of the nine individual panelists' perceptions as to whether or not each capability now exists in CIS. Values of five to nine in the yes column indicate a majority view that the capability is present; values of five to nine in the no column indicate a majority opinion that the capability is missing and should be added; values of five to nine in the partially column indicate a majority view that the capability is present in CIS, but needs improvement (e.g, consistency throughout system). Values of less than five in any column indicate divergent views by some panelists.

Current CIS capabilities and characteristics, based on panel perceptions, do not match those capabilities proposed by the panel. The existing system matches the proposed capabilities and characteristics completely in 4 of 27 areas, and partially matches in an additional 12 areas.

The panel considered the exceptional and essential aspects of CIS to include the CIS directory capability (e.g., SANSS) and numeric data bases. Characteristics most conspicuously missing included consistent mnemonics, consistent terminology and good educational and training materials.

The panel unaminously recommended that CIS be improved by adding the capabilities and characteristics which it currently lacks and improving the capabilities and characteristics now perceived to be only partially achieved.

3.3 Specific CIS Management Problems

Table 3-2 lists areas of decision-making by CIS managers and the panel's assessment of the adequacy of performance in each of eight areas. The panel's consensus was that all eight areas currently indicate inadequate performance with the areas of data base selection, data base implementation and management of public awareness requiring the most improvement. The panel identified possible causes of these performance inadequacies to be government operation of CIS, inadequate financial resources, previous management decisions, other factors or combinations of all of these factors. A combination of these factors was concluded by the panel to be the cause in each instance. Specific examples of the resulting problems appear in Table 3-3. These problems were cited as concrete reasons why CIS requires immediate improvement.

3.4 Steps Toward Becoming Self-Supporting

The panel agreed that CIS should not be required to become entirely self-supporting. The government has an on-going responsibility to the public to support underutilized, but valuable components of CIS. Specific actions, however, can and should be taken to make CIS more self-supporting. These actions fall into three categories: retrieve more costs, improve efficiency and increase the number of users. The specific recommendations of the panel by category are:

TABLE 3-1 COMPARISON OF PROPOSED WITH EXISTING CIS CAPABILITIES AND CHARACTERISTICS

Proposed Capabilities and Characteristics	Does Capability or Characteristic Presently Exist in CIS? Yes	No	Partially	Do Not Know
System Capabilities and Characteristics:				
1. Single, integrated system.	3	1	5	0
2. Ability to switch to outside system (e.g., CAS (Chemical Abstract Service) on-line).	0	4	5	0
3. Adequate directory to files and data (e.g., SANSS).	9	0	0	0
4. Substructure search capability.	3	0	5	1
5. Low-usage, but important, data bases provided.	8	0	1	0
6. Access to proprietary confidential data bases by government users (assuming security problems could be handled).	0	7	0	2
7. Adequate speed of operation.	0	1	8	0
8. Simultaneous searching capability.	0	6	3	0
Access Language Capabilities and Characteristics:				
1. Standardized, simplified, high-level, state-of-the-art language.	0	7	2	0
2. Data manipulation ability including ability to match with user-supplied data.	2	0	7	0
3. Multiple access points (exhaustive indexing).	4	0	5	0
4. Modeling abilities.	2	0	7	0
5. Ability to perform complex search combinations (Boolean operations).	6	0	3	0
6. Report writing ability.	0	6	3	0
7. Graphics capabilities.	0	3	6	0
8. Use of consistent mnemonics (e.g., for fields).	0	8	0	1

continued-

Table 3-1 - continued

Proposed Capabilities and Characteristics	Does Capability or Characteristic Presently Exist in CIS? Yes	No	Partially	Do Not Know
Data Base Capabilities and Characteristics:				
1. Kept updated.	0	6	3	0
2. Consistent use of terminology.	0	9	0	0
3. Actual data or information.	9	0	0	0
Data Capabilities and Characteristics:				
1. Use of individually evaluated data, wherever possible.	0	0	9	0
2. Incorporation of revised data whenever necessary.	0	7	2	0
3. Presence of chemical status data.	1	1	7	0
4. Ability to verify data (auditrail).	0	5	4	0
5. Limited to high quality data.	0	0	9	0
External Capabilities and Characteristics:				
1. Good user training program.	0	7	2	0
2. Good educational materials (e.g., manuals, sampler, newsletters).	0	8	1	0
3. Sufficient user support personnel.	0	0	9	0

TABLE 3-2 POSSIBLE CAUSES OF INADEQUATE CIS MANAGEMENT PERFORMANCE[a]

Decision-Making Area	Government Operation of CIS	Inadequate Financial Resources[b]	Previous Management Decisions[b]	Other Factors
1. Data base selection.	1	7	9	0
2. Data base implementation.	0	8	8	0
3. Data maintenance.	0	8	7	1
4. System standards for software, data documentation.	2	5	8	0
5. Quality control over data and software.	0	5	9	0
6. Policy-making including role of Management Board, EPA and other sponsoring agencies.	4	0	9	0
7. Management resources made available to CIS.	3	6	5	1
8. Public awareness of CIS.	9	6	9	3

(a) Values indicate number of votes by nine panelists for each possible cause.
(b) Panelists recognized that financial resources and management decisions are often interdependent.

Life Systems, Inc.

TABLE 3-3 EXAMPLES OF CIS PROBLEMS BY DECISION-MAKING AREA

Decision-Making Area	CIS Problems
Data base selection	• No data base selection policy. • Potential for quality data bases from government agencies not realized, for example, EPA, NIH, other agencies have many data bases not on CIS. • Inadequate resources to pull in other valuable data bases.
Data base implementation	• Implementation schedule not maintained. • Some data bases implemented with inadequate quality control.
Data base maintenance	• Data base developers do not adhere to announced maintenance schedules, therefore, erratic maintenance. • Irregular processing of updates.
System standards for software and data documentation	• Apparent nonexistent standards for software and data documentation. • Lack of software documentation has made change to a different computer environment nearly impossible. • Lack of data documentation has resulted in retention of files on system which users do not trust, contributing to a loss of credibility for the system. • Apparent nonexistent standards for system performance (e.g., speed of operation).
Quality control over data and software	• Lack of standards for data and software quality. • Lack of comprehensive quality control through processes such as independent peer review of data bases.
Policy-making including role of Management Board, EPA and other sponsoring agencies	• Literal lack of purpose. • Unlike name implies, Management Board has no management authority (i.e., advisory capacity only). • Existing CIS management has inadequate financial flexibility (e.g., sponsoring agencies limit funding to their sponsored files). • Lack of long-range planning due to limitations on contracts and funding.

continued-

Table 3-3 - continued

Decision-Making Area	CIS Problems
	• Interagency actions indicate lack of common goals (e.g., unwillingness to require other useful government data bases be implemented on CIS).
Management resources dedicated to CIS	• Management board concept is weak and faulty (i.e., labor energies not well spent without corresponding authority). • Management not sufficiently assertive over data base producers. • Inconsistent dedication of staff personnel.
Public awareness of CIS	• Inadequate use of available marketing services and techniques since end user market (e.g., chemists) not reached. • Lack of reinforced marketing (e.g., spasmodic interaction with users). • Poor training.

1. Retrieve More Costs.

   a. Eliminate free access, except for promotional costs during marketing.
   b. Add revenue generating products and services (e.g., off-line printouts, search services).
   c. Add other profitable data bases.
   d. Sell user manuals, training materials and training services.

2. Improve Efficiency.

   a. Reduce overhead.
   b. Eliminate data bases no longer of value (not based on use alone).
   c. Provide stronger guidance to data base developers.
   d. Reduce overlap in data bases.
   e. Match operating times to operating needs (e.g., reevaluate need for 24 hour accessibility).
   f. Use longer contract periods for contracted services.
   g. Improve cooperation between CIS management, user groups and data base developers by opening communication lines.

3. Increase the Number of Users.

   a. Eliminate start-up fee.
   b. Develop and initiate competitive marketing strategy.
   c. Improve documentation, training and other user services.
   d. Improve quality and currency of data.
   e. Standardize terminology throughout all data bases.
   f. Make system more user friendly (e.g., menu driven).

The panel recommended several other guidelines in the area of cost recovery from users. System access or operating costs, administration costs and user support costs should be recovered from users. On the other hand, continued government support should be provided for system development. For example, SANSS enhancement and data collection costs should be subsidized by the government, while SANSS usage costs should be charged to users. In establishing user fees, the following ideas should be considered as ways to benefit CIS while fairly distributing user costs: use of volume and off-peak-hour discounts, use of other price differentials based on ability to pay or demand on the data base, charging according to fair market value of data bases, direct charges for time on-line and for output, in addition to connect time.

3.5 Data Bases or Files Which Could Now Be Self-Supporting In the Private Sector

The panel agreed that private data base vendors with sufficient resources and the capability for generating sophisticated software could easily implement and provide all of the existing individual CIS files. Based on the panel's judgment and the limited background information (there was a lack of complete economic information) provided by EPA, the following existing CIS files could be self-supporting through private data base vendors:

1. Structure and Nomenclature Search System (SANSS).
2. Registry of Toxic Effects of Chemical Substances (RTECS).
3. Oil and Hazardous Material Technical Assistance Data System (OHMTADS).
4. Mass Spectral Search System (MSSS).
5. Scientific Parameters in Health and the Environment, Retrieval and Estimation (SPHERE) (very limited data for judgement).
6. Toxic Substances Control Act Plant and Production (TSCAPP) (EPA must maintain private file for confidential information).
7. Clinical Toxicology of Commercial Products (CTCP).
8. Federal Register Search System (FRSS).

The panel, however, noted that the impact of moving some CIS files to private vendor(s) and not having them available as an integrated system would be detrimental to CIS. Those data bases which might be most profitable, and therefore, most attractive to a private vendor, might be the same CIS resources most able to help the system be self-supporting.

The panel recommended that if CIS is improved to meet the specifications developed by this panel, these data bases should remain part of CIS. This recommendation does not preclude concurrent availability in the private sector which might be a source of additional revenue to CIS.

3.6 Government Responsibility

The panel recommended that the following statement be adopted as a management guideline: Given that availability of chemical information is in the general public interest, the Federal government has the responsibility to provide data or information not otherwise available at a reasonable cost and reasonable response time in a publicly available system.

3.7 Appropriate CIS Management Structure

3.7.1 Responsibilities

In order to clarify their evaluation of the appropriate management structure, the panel defined the individual responsibilities of the management and the implementing organization toward CIS. Specific responsibilities of the management organization should be to:

1. Ensure that the system is meeting the defined needs.
2. Ensure accessible and affordable chemical data to all users.
3. Ensure appropriate content of the data bases and system.
4. Fund, as necessary, major development of systems capabilities.
5. Furnish appropriate government data bases.
6. Ensure that underutilized, but valuable data bases are available/provided in a system.
7. Guide and enhance SANSS.

Specific responsibilities of the organization that operates and maintains CIS should be to:

1. Install and operate data bases.
2. Ensure that supplied data bases are current, of high quality and that the formats have the necessary access points and are properly indexed.
3. Maximize cost effectiveness to design and implement systems enhancement.
4. Standardize the system.
5. Develop user friendly access.
6. Conduct financial management.
7. Perform appropriate marketing.
8. Respond to all user input.
9. Carry out user support services.
10. Advise government on planning aspects.
11. Advise government on system enhancement.

3.7.2 Advantages and Disadvantages for Potential Implementing Organizations

After review of existing management performance, it became clear to the panel that an inter-agency group is not capable of achieving the proposed purpose of CIS for the following reasons:

1. Government groups are not experts in data base vending or marketing (e.g., Remote Console (Information retrieval system operated by the Department of Energy (DOE)) (DOE/RECON) versus DIALOG).
2. Inter-agency cooperation is insufficient.
3. Uncertainty in year-to-year commitments of funds and personnel from participating agencies hampers long-term planning.
4. Financial management of the system (i.e., billing and cost recovery) is inefficient.
5. Poor responsiveness to non-government users is the norm.

The panel further recognized that individual agencies are not appropriate for implementing CIS, with the exception of the National Library of Medicine (NLM).

The panel agreed that the NLM offers advantages as an organization to operate and maintain CIS:

1. Single government agency.
2. Existing hardware.
3. Appropriate Experience.
4. Currently operates related information systems.
5. History of commitment to automated data bases.
6. Commitment to state-of-the-art data bases.
7. Good at frequent updating.
8. Large diverse user groups.
9. Good reputation for training.
10. Financial management generally sound.
11. Above average marketing ability for a government agency.

In contrast, the disadvantages of NLM as an implementing organization are few:

1. Lack of interest in promoting CIS as a system.
2. Encumbered by government red tape.
3. Not as good at marketing as private sector.
4. Poor billing procedures.

In seeking the appropriate type of organization to operate and maintain CIS, the panel considered the following additional options:

1. A non-profit organization established by the Federal Government for the purpose of implementing CIS.
2. An existing private vendor under contract (profit-making organization).
3. A professional society (e.g., ACS, FASEB).
4. A privately operated government laboratory (e.g., Oak Ridge, Brookhaven).
5. A private non-profit research institute (e.g., Battelle, Stanford Research Institute (SRI)).

The panel developed a list of advantages and disadvantages for each option. Associated with the government-established non-profit organization option are the following advantages:

1. Dedication to one system.
2. Non-profit status.
3. Direct contact and responsiveness to agencies supplying data bases.
4. Backed by Federal Government.
5. Competitive in marketplace.

This option, however, has the following disadvantages:

1. Act of Congress required to establish.
2. Creation of new organization necessary (e.g., assembly of experienced personnel).
3. Critical mass problems and high overhead due to small size.
4. Tendency to not be responsive to users.
5. Tendency to become governmental.

A private vendor would offer the following advantages:

1. Large size (possibly).
2. Experience.
3. Existing hardware.
4. Operating efficiencies.
5. Responsive to users.
6. Sound financial management.
7. Financial incentive for improvement.
8. Marketing expertise and incentives.
9. Tendency to eliminate poor inefficient underutilized data bases.

Disadvantages of private vendors are:

1. Lack of interest in CIS as a system.
2. Requirement for profit.

3. Tendency to neglect responsibility for updating and enhancing data bases and system.
4. Tendency to not be as responsive to data base suppliers (unless data base vendor is funding).
5. Tendency to eliminate valuable underutilized data bases.

The professional society option offers the following advantages:

1. Large size vending operation.
2. Experience.
3. Existing system.
4. Responsive to users.
5. Established user group.
6. Maximized service to professionals.
7. Sound financial management.
8. Financial incentive for improvement.
9. Tendency to keep valuable underutilized data bases.
10. Tendency to eliminate poor inefficient underutilized data bases.

The disadvantages of the professional society are:

1. Lack of interest in CIS as a system.
2. Lack of marketing expertise (relative to large private data base vendors).
3. Tendency to not be as responsive to data base suppliers.

It was perceived that the advantages and disadvantages of a privately operated government laboratory would be similar to the private non-profit corporation, but would not require a new Congressional Statute to become established. The private non-profit research institute option was considered to be similar to the professional society in terms of advantages and disadvantages.

3.7.3 Recommendation for Implementing Organization

After weighing all the advantages and disadvantages, the panel concluded that there were three major candidates for an appropriate implementing organization for CIS:

1. NLM
2. A professional society
3. A private vendor

As a secondary position, the panel felt that a privately operated government laboratory or a private research institute could also be viable implementing organizations. The panel believed the private non-profit corporation established by the government was not a realistic candidate due to its requirement to be established by an Act of Congress.

3.7.4 Appropriate Management Organization

The panel recognized that the government is clearly responsible for assuring the existence of a CIS that meets the needs of the government and other users. Therefore, they recommended the following:

1. The responsible group within the government should be EPA, NLM, or CEQ. This ranking indicates the order of choice by the panel. The group within EPA, if EPA is chosen, should be selected by EPA or by the Ad Hoc Government Panel.

2. In order to assist the lead group an interagency advisory group of government users and data base providers should be established to review CIS and its data bases with respect to government needs.

3. Each agency sponsoring data bases on CIS should pay for the development of those data bases it provides.

3.7.5 Summary

Therefore, in order to better achieve the purpose of CIS, the panel recommended adoption of a more appropriate management concept consisting of a responsible management role for the government and a separate implementation role. For the responsible management role within the government the panel recommended EPA, NLM or CEQ in that order of preference. The panel identified NLM, a professional society, or a private data base vendor, as viable candidates for the implementation role. As a second choice, the panel identified a government laboratory or a private research institute as additional candidates.

APPENDIX 1

TR-576-42B

LIST OF PARTICIPANTS

REVIEW OF CHEMICAL INFORMATION SYSTEM SYMPOSIUM

Panelists:

Dr. William Bailey
Department of Chemistry
University of Maryland
College Park, MD 20742

Dr. Edward Bartkus(a)
Executive Information Management
507 Falkirk Road
Wilmington, DE 19803

Dr. Trudi Bellardo(a)
College of Library and
Information Science
University of Kentucky
Lexington, KY 40506-0027

Dr. James Brower(b)
CACPH, Bldg. 535A
Brookhaven National Laboratory
Upton, Long Island, NY 11973

Ms. Linda Greer
Environmental Defense Fund
1525 18th Street, NW
Washington, DC 20036

Dr. Warren Muir
4158 South 36th Street
Arlington, VA 22206

Mrs. Edna Paulson
National Research Council
Toxicology Research Center
2101 Constitution Avenue
Washington, DC 20418

Dr. Alberta Ross
Radiation Laboratory
University of Notre Dame
Notre Dame, IN 46556

Mrs. Lynda Wiseman
Celanese Corporation
1211 Avenue of the Americas
New York, NY 10036

continued-

(a) Designated alternate spokesperson for the panel.
(b) Designated spokesperson for the panel.

Other Participants:

Mr. Larry Dusold
U.S. Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857

Dr. Sherman P. Fivozinski
National Bureau of Standards
Office of Standard Reference Data
Physics Bldg. Room 318
Washington, DC 20243

Dr. Lewis H. Gevantman
11608 Toulone Drive
Potomac, MD 20854

Dr. George W.A. Milne
National Cancer Institute
9000 Rockville Pike
Bethesda, MD 20205

Ms. Katherine Noble
Computer Sciences Corporation
P.O. Box 227
Falls Church, VA

Dr. Pat Shannon
U.S. Consumer Product Safety Commission
Washington, DC 20207

Dr. Louis R. Sibal
9000 Rockville Pike Bldg. 1
Room 314
Bethesda, MD 20205

EPA

Tony Jover
Sarah Kadec
Mary Lou Melley
Howard Messner

ICAIR, Life Systems, Inc.

Ms. Erin Dayl
Mr. KAI Dozier
Mr. Jeffrey Heaton
Ms. Patricia Lavan

APPENDIX 2

BACKGROUND REFERENCES SUPPLIED TO PANELISTS

Altman PL, Fisher KD (ed.). 1982. A user assessment of the Chemical Information System. Bethesda, MD: Life Sciences Research Office, Federal of American Societies for Experimental Biology. Prepared for American Management Systems, Inc. Subcontract No. 2881-1, DTO-5146-50. EPA Contact No. 68-01-5146.

Anonymous. 1982. Agreement for operation of the Chemical Information System (CIS).

Anonymous. Chemical information files and programs, data sheets: ARTHUR, CCRIS, CESARS, CHEMLAB, CNMR, CRYST, CTCP, FRSS, IRSS, MLAB/CLAB, MSSS, OHMTADS, RTECS, THERMO, XTAL.

Anonymous. 1984. Chemical Information System usage data for all users (September 1983 - February 1984).

Anonymous. 1984. Chemical Information System usage data for EPA users (September 1983 - February 1984).

Anonymous. 1984. NIH/EPA Chemical Information System FY 1982/FY 1983/FY 1984 expenditure summary comparison as of March 6, 1984.

Anonymous. 1982. NIH/EPA Chemical Information System. Status Report No. 15, June 1982.

CIS User Support Group. 1983. CIS search sampler, July 1983. Falls Church, VA: Computer Sciences Corporation.

Council on Environmental Quality. 1978. The feasibility of a standard chemical classification system and a standard chemical substances information system: a report to the Congress prepared pursuant to section 25(b) of the Toxic Substances Control Act of 1976 (15 U.S.C 2601). Washington, DC: U.S. Government Printing Office, Stock No. 041-011-00039-4.

CRC Systems, Inc. 1983. EPA Chemical Information System overview and requirements. Fairfax, VA: CRC Systems, Inc. EPA Contract No. 68-01-6746, D.O. No. 8.

Gevantman LH. 1984. Status of data files in the Chemical Information System (CIS). Memorandum to A. Jover.

Heller SR, Potenzone Jr. R, Milne GWA, Fisk C. 1981. Computers in analytical chemistry. Trends in Analytical Chemistry 1(3):62-65.

Heller SR. 1984. Unique capabilities of CIS. Memorandum to S. Kadec. January 16.

ICAIR, Life Systems, Inc. 1984. Chemical Information System (CIS) user survey tabulation. Cleveland, OH: Life Systems, Inc. EPA Contract No. 68-02-4038, TR-576-44.

Meschel SV. 1984. Numeric databases in the sciences. Online review 8(1):77-103.

Milne GWA. 1982. Development of a Chemical Information System. J. Assoc. Off. Anal. Chem. 65(5):1249-1258.

Milne GWA, Fisk CL, Heller SR, Potenzone Jr. R. 1982. Environmental uses of the NIH-EPA Chemical Information System. Science 215:371-375.

USEPA. 1984. EPA journal. Vol. 10, No. 3. Washington, DC: U.S. Environmental Protection Agency.

APPENDIX 3

TR-576-43B

AGENDA[a]

Review of Chemical Information System Symposium

June 11-13, 1984
Holiday Inn
Bethesda, MD

Time	Agenda Item	Individual
Monday, June 11, 1984		
8:50 a.m.	Welcome	J. Heaton
9:00 a.m.	Opening and Charge	H. Messner
9:15 a.m.	CIS Overview	K. Noble
9:45 a.m.	Approach to Symposium Goals	K. Dozier
10:15 a.m.	Break	
10:30 a.m.	Session No. 1: Overview	K. Dozier
12:00 Noon	Luncheon for Participants	
1:00 p.m.	Hardware Demonstration	K. Noble
2:00 p.m.	Session No. 2: Systems Specifics--Capabilities	K. Dozier
3:45 p.m.	Break	
4:00 p.m.	Open Discussion	K. Dozier
5:00 p.m.	Adjourn	
7:00 p.m.	Terminals Available for Use	
9:00 p.m.	Terminals Closed	

continued-

(a) Agenda was modified during symposium as necessary to accomplish overall objectives.

Appendix 3 - continued

Time	Agenda Item	Individual
Tuesday, June 12, 1984		
8:00 a.m.	Question Period with CIS Management Board	L. Dusold
9:00 a.m.	Session No. 2 continued: System Specifics--Data	K. Dozier
10:30 a.m.	Break	
10:45 a.m.	Session No. 2 continued: System Specifics--Technical Areas and Cost	K. Dozier
12:15 p.m.	Break	
1:30 p.m.	Use of Terminals	K. Noble
2:30 p.m.	Session No. 2 continued: System Specifics--Cost continued and Management	K. Dozier
3:45 p.m.	Break	
4:00 p.m.	Open Discussion	K. Dozier
5:00 p.m.	Adjourn	
7:00 p.m.	Terminals Available for Use	
9:00 p.m.	Terminals Closed	
Wednesday, June 13, 1984		
8:00 a.m.	Question Period with CIS Management Board	L. Dusold
9:00 a.m.	Session No. 2 continued: System Specifics--Management continued	K. Dozier
10:30 a.m.	Break	
10:45 a.m.	Session No. 3: Management and Accessibility	K. Dozier
12:15 p.m.	Break	

continued-

Appendix 3 - continued

Time	Agenda Item	Individual
Wednesday, June 13, 1984 - continued		
1:30 p.m.	Use of Terminals	K. Noble
2:00 p.m.	Session No. 3 continued: Management and Accessibility	K. Dozier
3:15 p.m.	Break	
3:30 p.m.	Summary	K. Dozier
4:00 p.m.	Final Recommendations	K. Dozier
5:00 p.m.	Close	K. Dozier J. Heaton